DOCUMENT RESUME

ED 146 218 TM 006 613

AUTHOR DeVito, Pasquale J.
TITLE The Use of Commonality Analysis in Educational Research.
PUB DATE May 76
NOTE 18p.; Paper presented at the Annual Meeting of the New England Educational Research Association (Provincetown, Massachusetts, May 5-8, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Cognitive Ability; *Compensatory Education Programs; *Componential Analysis; Disadvantaged Youth; Educational Background; Intermediate Grades; Parent Participation; *Predictor Variables; *Reading Achievement; Reading Comprehension; *Statistical Analysis; Student Characteristics; Vocabulary Skills
IDENTIFIERS *Commonality Analysis; Partitioning Procedures

ABSTRACT

Commonality analysis is a statistical technique used within the context of a study to examine school effects among disadvantaged students. The research investigated the unique and common contributions of background, mental ability, program and parental involvement variable sets to the reading vocabulary and comprehension of students participating in a compensatory education program funded under Title I. The major advantage of commonality analysis is that both the unique and common contributions of the variables or sets of variables to the variance of the dependent variable can be identified. Several disadvantages of this method are also discussed. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146218

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE USE OF COMMONALITY ANALYSIS IN EDUCATIONAL RESEARCH

Pasquale J. DeVito
Rhode Island Department of Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Pasquale J. DeVito

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

A paper presented at the Annual Meeting of the New England Educational Research Association.

Provincetown, Mass.
May 5-8, 1976

TM006 613

INTRODUCTION

The main purpose of this paper is to explain commonality analysis as a technique, its advantages and disadvantages, within the context of a study to examine school effects among disadvantaged students. The research attempted to investigate the unique and common contributions of background, mental ability, program and parental involvement variable sets to the reading vocabulary and comprehension of students participating in compensatory education.

Sample and Data Collection

The sample used in this study included 877 students in grades 4-6 enrolled in Title I remedial or corrective reading programs in the state of Rhode Island in FY 1973-74 and for whom there were appropriate grade level pre-test and post-test Gates-MacGinitie Reading Tests vocabulary and comprehension scores.

Data used in this study were collected as part of the usual Rhode Island Department of Education's Office of Compensatory Programs. Throughout the program year information was collected via four state reporting forms and the return rate was near 100 percent.

Variables Included in the Study

The following variable sets were identified:

Set 1: Background Variables

1) Type of community
2) Sex
3) Ethnic Group
4) Prior years in Title I reading programs
5) Type of school attended
6) Number of times retained in a grade

Set 2: Mental Ability Variables

1) IQ
2) Pre-test reading scores

Set 3: Program Variables

1) Pupil-teacher ratio
2) Per-pupil expenditure
3) Length of project
4) Number of days student was absent
5) Minimum amount of individual instruction per student per week
6) Size of instructional group for students
7) Number of children serviced per week
8) Amount of scheduled preparation time per week with regular teacher to discuss students
9) Whether materials were available at each child's instructional level

10) Whether materials were available on time for project start
11) Whether teachers selected materials
12) Amount of time spent by teachers per week develooping their own materials
13) Whether pre-service or in-service training activities were held for staff

Set 4: Parental Involvement Variables

1) How often parents were responsible for working with students at home
2) Whether each parent was seen at least once during project year
3) Whether Parents Advisory Committee made recommendations on expenditures of Title I funds
4) Whether Parents Advisory Committee participated in the development of Title I applications
5) Whether Parents Advisory Committee reviewed Title I applications
6) Whether Parents Advisory Committee made recommendations on improvement of Title I programs
7) Whether Parents Advisory Committee participated in Title I program evaluation

Treatment of the Data

Commonality analysis (also called elements or component analysis) is a method of analyzing the variance of a dependent variables into common and unique to identify the relative influences of independent variables or sets of variables. It is an attempt to understand the relative predictive power of the regressor variables, both individually and in combination. The squared multiple correlation is broken up into elements assigned to each individual regressor or set and to each possible combination of regressors or sets. The elements have the property that the appropiate sums not only add to squared multiple correlations with all regressors, but also to the squared multiple correlation of any subset of variables, including the simple correlations.

Mood (1971) presented an example, using two sets of variables, of how unique and common contributions are found.

"Let us suppose that the first m of the x's are intended to be indicators of X_1 and refer to them as the W set of x's; let us lump all the other n-m's into another set and refer to it simply as the Y set. We are going to partition the variance attributable to the regression of A on the x's into three parts - rahter we shall use the multiple correlation instead of the variance. We first calculate three regressions:

A on the W set of x's only
A on the Y set of x's only
A on the whole set of x's

and let us suppose that the first removes 20 percent and the raw variance of A, the second removes 55 percent, and the third removes 60 percent. Now we divide the 60 percent removed by the whole Set (W + Y) into three parts:

a part uniquely associated with W, 5%
a part uniquely associated with Y, 40%
a part that may be associated with either W or Y 15%

The part uniquely associated with the W set is calculated by subtracting the proportion removed by the Total (W + Y) Set. The reason for attributing this 5 percent uniquely to W is simply that the x's in the Y set removed 55 percent of the 60 percent removed by the total; on adding the W set to the Y set we remove only an extra 5 percent so that it is the part that must be uniquely associated with W. Similarly, the W set alone removes 20 percent; on adding the Y set to it we remove only an additional 40 percent so that it is the part that must be uniquely associated with Y.

Table I

	W	Y
Unique to W	5%	
Unique to Y		40%
Common to W and Y	15%	15%
Totals	20%	55%

Finally, the part that may be associated with either W or Y is calculated by subtracting the two unique parts from the whole (60% - 5% - 40%) (Mood, 1971, p. 194-195.)"

Advantages of the Method

The major advantage of the technique is that the researcher can identify both the unique and common contributions of the variables or sets of variables to the variance of the dependent variable. This capability offers distinct advantages over more frequently used, traditional, types of analyses like analysis of covariance or step-wise multiple regression.

Elashoff (1969) has stated that analysis of covariance should be considered when the investigator believes that some outside variable will have a large distorting effect on the results and when the assumptions of normality of data and random assignment of subjects to treatment are met. She stated that ANCOVA is widely used to "adjust" criterion scores such as achievement for the effects of a covariate such as ability in order to compare several treatments. However, in school effects studies like this one, as well as in other educational research, the investigator may be more interested in looking at all effects and contributions of variables than in controlling statistically for the effect of a variable.

Similarily step-wise multiple regression is an often used approach to handling data. It does allow the researcher to identify the unique contribution of each variable to the variance of the dependent variable by determining the increase in explained variance by adding each variable to the regression equation. It is impossible, however, to determine the joint or common contributions of the sets or variables through the use of step-wise multiple regression procedures. Often an investigator can gain insight into educational models by looking at the predictive power variables share with one another.

In addition to these benefits, Mood (1971) suggested that not only individual variables but also sets of variables representing some factor could be used as independent variables in the analysis.

Disadvantages

Like any relatively new analysis technique, there are some problems and disadvantages in using this method.

Several areas of concern should be noted in relation to commonality analysis. The first deals with the difficulties encountered in testing for significance. Mood (1971) stated that one could make the usual F test of significance for unique parts to determine whether additional regression terms have contributed significantly to the regression. One cannot, however, test the common parts for significance. This concern is not a major one here since this large sample study is more interested in unique and common contributions of the factors to the dependent variables than in statistical significance.

A second concern deals with the interaction of sets of variables. Tatsuoka (1973) stated that the relationships between the joint contributions of sets of variables should not be confused with the interaction of these sets. However, if one were interested in the interaction, the product term method could be used. Kerlinger and Pedhazur (1973) cited an example in the two variable case. Assuming one had two variables, X_1 and X_2, the values of these two independent variables could be multiplied over all cases to create a third variable, X_1X_2. This variable is then entered into the regression equation as another variable, and, if there is a significant interaction between the variables in their effect on the dependent variable, it will be evident in the significance test. The analysis used in this study, however, was designed to investigate the unique and common contributions of the sets of variables on the dependent variable and not the interaction of these factors.

A third concern is that some of the commonalities can have negative signs. Kerlinger and Pedhazur (1973) stated that negative commonalities

can be obtained in situations where one of the variables is a suppressor, or when correlations among independent variables are negative. Negative proportions of shared variance among variables can be difficult to interpret.

A fourth concern centers around the number of variables or sets to be included in analysis. Unique contributions are presented for each dependent variable or sets as well as common contributions presented for every possible combination. If a great number of variables or sets are used, interpretation may become difficult and unwielding.

RESULTS

The data were analyzed by using the partitioning of variance technique in multiple regression to determine the unique and joint contributions of four sets of variables in the reading achievement of compensatory education students in Rhode Island. Separate analyses were conducted for vocabulary and comprehension scores.

Four sets of variables were included in the analyses. The background set included six variables. Two variables in the set - type of community and ethnic group - necessitated the construction of dummy variable coding for these indicators. The mental ability set included two variables. The reading program set included thirteen variables related to instructional and program elements. The parental involvement set included six variables in the analysis. Seven variables were initially intended to be used; however, preliminary analyses indicated a high relationship between two - participation of Parents Advisory Committee in the development of Title I applications and review of Title I applications by Parents Advisory Committee - about +.80, so an additional variable combining these two was constructed.

Vocabulary Analysis

The first analysis was performed using the four sets specified above with post-test vocabulary standard scores as the dependent variable. The

total amount of explained variance accounted for by the four sets of variables was sixty percent.

Background Set -

The unique contribution of the set of six background variables to the variance of vocabulary scores was about seven percent. In addition, the overlap variance, i.e., that variance shared jointly with other sets was about ten percent.

Mental Ability Set -

The unique contribution of the set of two mental ability variables to the variance of vocabulary scores was about thirteen percent. The overlap variance associated with this set was about nine percent.

Program Set -

The unique contribution of the set of thirteen program-related variables to the variance of vocabulary scores was about nine percent. In addition, the overlap variance for this set equaled about five percent.

Parental Involvement Set -

The unique contribution of the set of six parental involvement variables to the variance in vocabulary scores was about nine percent. Also, the overlap variance associated with this set was seven percent.

Only unique and overlap contributions have been mentioned here; however, the complete set of commonality coefficients are presented in Table 1.

Comprehension Analysis

The second analysis was performed using the four sets of variables specified and post-test comprehension standard scores as the dependent variable. The total amount of explained variance accounted for by the four sets of variables was about forty-six percent.

Background -

The unique contribution of the set of six background variables to the variance of comprehension scores was about seven percent, while the overlap variance associated with the background set was about four percent.

Mental Ability Set -

The unique contribution of the set of two mental ability variables to the variance in comprehension scores was about twenty-five percent while overlap variance was about four percent.

Program Set -

The unique contribution of the set of thirteen program-related variables to the variance in comprehension scores was about five percent, while the overlap variance was slight, about .4 percent.

Parental Involvement Set -

The unique contribution of the set of six background variables to the variance in comprehension scores was about four percent. The overlap variance was about three percent.

While unique and overlap contributions have been presented here, the complete set of commonality coefficients for the comprehension analysis are shown in Table 2.

NEGATIVE COMMONALITIES

When partitioning of variance technique is used, there exists the possibility of obtaining negative commonalities, that is, to obtain negative proportions of shared variance. Beaton (1974) stated that the unique elements must be non-negative but the common parts may be either positive, negative or zero. He also mentioned that negative commonalities are not usually

TABLE 1

PERCENTAGES OF EXPLAINED VARIANCE OF FOUR SETS

OF VARIABLES ON READING VOCABULARY

AFTER PARTITIONING

	Sets of Variables			
	1	2	3	4
Unique to Set 1, (Background)	.0689			
Unique to Set 2, (Mental Ability)		.1313		
Unique to Set 3, (Program)			.0867	
Unique to Set 4, (Parental Involvement)				.0943
Common to Sets 1 and 2	.0398	.0398		
Common to Sets 1 and 3	.0842		.0842	
Common to Sets 2 and 3		.0228	.0228	
Common to Sets 1 and 4	.0580			.0580
Common to Sets 2 and 4		.0563		.0563
Common to Sets 3 and 4			.0900	.0900
Common to Sets 1, 2, and 3	.0031	.0031	.0031	
Common to Sets 1, 3, and 4	-.1003		-.1003	-.1003
Common to Sets 2, 3, and 4		-.0466	-.0466	-.0466
Common to Sets 1, 2, and 4	.0150	.0150		.0150
Common to Sets 1, 2, 3, and 4	-.0036	-.0036	-.0036	-.0036
	.1651	.2182	.1362	.1631

TABLE 2

PERCENTAGES OF EXPLAINED VARIANCE OF FOUR SETS OF VARIABLES ON READING COMPREHENSION AFTER PARTITIONING

	Sets of Variables			
	1	2	3	4
Unique to Set 1, (Background)	.0682			
Unique to Set 2, (Mental Ability)		.2511		
Unique to Set 3, (Program)			.0547	
Unique to Set 4, (Parental Involvement)				.0436
Common to Sets 1 and 2	.0147	.0147		
Common to Sets 1 and 3	-.0073		-.0073	
Common to Sets 2 and 3		.0042	.0042	
Common to Sets 1 and 4	-.0081			-.0081
Common to Sets 2 and 4		.0242		.0242
Common to Sets 3 and 4			-.0062	-.0062
Common to Sets 1, 2 and 3	-.0059	-.0059	-.0059	
Common to Sets 1, 3 and 4	.0156		.0156	.0156
Common to Sets 2, 3 and 4		-.0254	-.0254	-.0254
Common to Sets 1, 2 and 4	.0020	.0020		.0020
Common to Sets 1, 2, 3 and 4	.0289	.0289	.0289	.0289
	.1081	.2944	.0586	.0752

found in educational research. This statement may be a bit premature since partitioning of variance technique has been used in relatively few educational studies and has only recently been identified as a promising method in educational research.

The results of this study indicated some negative commonalities, most notably a - .1003 value for the joint contributions of the background, program and parental involvement sets in the vocabulary analysis. The interpretation of negative commonalities is not clear since the methodology is still in the developing stages.

Negative commonalities are clearly possible in partitioning of variance. A hypothetical example should make this evident. Assumethe two variable case, where Variable A and Variable B are use to predict the criterion Variable C, with the following squared multiple correlations: $R^2_{AC} = .50$, $R^2_{BC} = .00$, $R^2_{AB} = .40$, and $R^2_{ABC} = .60$. To determine the unique contribution of variable A, the following formula could be used:

$$U_A = - R^2_{BC} + R^2_{ABC} = - .00 + .60 = .60.$$

To determine the unique contributions of variable B, the following formula could be used:

$$U_B = R^2_{AC} + R^2_{ABC} = -.50 + .60 = .10.$$

The common contribution of variables A and B could be determined by the following formula:

$$C_{AB} = R^2_{ABC} - U_A - U_B = .60 - .60 - .10 = -.10.$$

The -.10 value for C_{AB} represents a negative proportion of shared variance.

Several authors (Newman and Newman, 1975, Kerlinger and Pedhazur, 1973) acknowledge this as a conceptual problem yet offer little direction or explanation to solve the fifficulty. Other writers supply more direction and information.

Veldman (1975a) suggested that in a situation like the one presented above a negative connonality results from a suppressor variable. A suppressor variable is related to another predictor variable yet unrelated to the criterion. In this way the variable suppresses the variance in another predictor which is unrelated to the criterion. The prediction of the criterion is increased by the inclusion of a suppressor variable into the regression equation.

The correlation matrix of variables was examined for indications of suppressor variables. Several instances of this type of relationship were found.

Beaton (1973) and Veldman (1975b) have also suggested that negative commonalities can occur when correlations between independent variables or sets of variables are negative. In this situation one variable or set actually confounds the predictive power. Beaton (1973) gives an example of a relationship of this type:

> "Both weight and speed are important to success as a professional football player and each would be moderately correlated with a measure of success in football. Weight and speed are presumable negatively negatively correlated and would have a negative commonality in predicting success in football. If both weight and speed are known, one would expect to make a much better prediction of success using both variables to select fast, heavy men rather than just selecting the fastest regardless of weight or heaviest regardless of speed. Thus the negative commonality indicates that explanatory power of either is greater when the other is used (Beaton, 1973, p. 22)."

In order to shed some light on possible negative correlation between variables within sets, the correlation matrix was again examined. Examples of this type of relationship were found to exist.

For purposes of interpreting the negative commonalities found in this study several statements and cautions should be made.

1) The coefficients presented in Tables 1 and 2 are shown to the precision of four decimal points. This was done to indicate complete results and in the event that this type of precision would be useful to readers of this study; however, it is certainly defensible to round off several of the coefficients.

2) If this is done, many of the negative commonalities presented in the tables become essentially zero. After rounding, no negative commonalities appea for any second order combination, only in the third order joint contributions.

3) Given the nature of the variable sets as well as some indications from the correlation matrix that negative correlations between some variables exist, it is the opinion of this writer that the negative commonalities are more likely to be due to negative correlations between sets than in suppressors. In many cases variables in each set of predictors were related positively to the dependent variables but negative correlations between variables in different sets existed.

4) When interpreting tables, Veldman (1975b) suggested that when negative commonalities are obtained, the independent contributions of the sets or variables involved are collectively overestimated.

5) Perhaps as educational models become better defined, the occurrence of negative commonalities will diminish. However, as commonality analysis is increasingly utilized, further research and guidelines on interpretation of these scores should be developed.

In conclusion, this investigation utilized a newly developed methodology, commonality analysis, in a school effects study. The technique

provided several advantages over more traditional types of analysis and proved highly satisfactory in the study. Negative commonalities were encountered in the analysis and attempts were made to adequately interpret the values. With some further developmental work the method should prove to be of benefit in future educational research ventures.

REFERENCES

Beaton, Albert, Commonality, Princeton, N.J., (mineo), 1973.

Elashoff, Janet B., "Analysis of Covariance: A Delicate Instrument", American Educational Research Journal, Vol. 6, No. 3, 1969, 383-403.

Kerlinger, Fred H. and Pedhazur, Elazar J., Multiple Regression in Behavioral Research, New York, Holt, Reinhart and Winston, 1973.

Mood, Alexander, "Partitioning Variance in Multiple Regression Analyses as a Tool for Developing Learning Models, "American Educational Research Journal, 1971, Vol. 8, No. 2, 191-202.

Newman, Isadore and Newman, Carole, A Discussion of Component Analysis: Its Intended Purpose, Strengths and Weaknesses, paper presented at Annual Convention of the American Educational Research Association, Washinton, D.C., 1975.

Veldman, Donald J., COMMAP-A Computer Program for Commonality Analysis, (mimeo), Austin, Texas, 1975.

Veldman, Donald, personal communication, letter, June 1975.